Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Real-time Vision-based Multiple Vehicle Detection and Tracking for Nighttime Traffic Surveillance

Yen-Lin Chen[1], Bing-Fei Wu[2,*], and Chung-Jui Fan[2]
[1]Dept. of Computer Science and Information Engineering, National Taipei University of Technology, Taipei, Taiwan
e-mail: ylchen@csie.ntut.edu.tw
[2]Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan
e-mail: *bwu@cssp.cn.nctu.edu.tw

***Abstract*—This study presents an effective system for detecting and tracking moving vehicles in nighttime traffic scene for traffic surveillance. The proposed method identifies vehicles based on detecting and locating vehicle headlights and taillights by using the techniques of image segmentation and pattern analysis. First, to effectively extract bright objects of interest, a fast bright-object segmentation process based on automatic multilevel histogram thresholding is applied on the nighttime road-scene images. This automatic multilevel thresholding approach can provide robustness and adaptability for the detection system to be operated well under various illumination conditions at night. The extracted bright objects are processed by a spatial clustering and tracking procedure by locating and analyzing the spatial and temporal features of vehicle light patterns, and then identifying and classifying the moving cars and motorbikes in the traffic scenes. Experimental results demonstrate that the proposed approach is feasible and effective for vehicle detection and identification in various nighttime environments for traffic surveillance.**

***Keywords*—Intelligent transportation systems, vehicle detection, vehicle tracking, nighttime surveillance, traffic surveillance.**

I. INTRODUCTION

Detecting and recognizing moving vehicles in traffic scenes for traffic surveillance, traffic control, and road traffic information systems is an emerging research area for Intelligent Transportation Systems. The information of moving vehicles may be obtained from the sensor such as loop detectors, slit sensors, or cameras. Among the above-mentioned sensors, camera-based systems can provide much more information for traffic analysis, such as traffic flow, vehicle classification, and vehicle speed.

Due to the progress on the reducing cost and growing computing power of the hardware, the vision-based technologies has become the popular solutions for traffic surveillance and control systems. To-date, many researchers have developed valuable techniques for detecting and recognizing vehicles and obstacles from images of traffic scenes [1]- [10]. By monitoring the illumination change in some pre-specified detection regions, the techniques based on virtual slit [1] or virtual loop detectors [2] can rapidly detect the moving vehicles when passing these regions. However, such methods are limited to obtain the number of passing vehicles passing through the given detection regions, and are difficult to apply on vehicle classification, vehicle speed detection, and vehicle motion analysis. To more efficiently obtain the traffic information of moving vehicles, frame-differencing based techniques are applied for segmenting the moving vehicles from motionless background scenes by change detection or some statistical models. In [3][4], spatial-temporal difference features are applied for segmenting the moving vehicles, while the methods in [5][6][7][9] utilize the background subtraction based techniques for extracting moving vehicles. These methods can efficiently apply on the daytime traffic scenes with stationary and unchanged light conditions.

However, under conditions of bad-illuminated at nighttime, the background scenes are substantially affected by lighting effect of moving vehicles, so that those obvious cues of background models which are effective for vehicle detection during daytime become invalid. Thus, most of the above-mentioned frame-differencing based techniques may not work well under such nighttime traffic environments. At night, as well as under darkly illuminated conditions in general, the only salient features of moving vehicles are their headlights and taillights.

To detect salient objects for nighttime traffic surveillance, Beymer *et al.* [8] presented a feature-based technique by extracting and tracking the set of corner features of moving vehicles instead of their entire regions, and can work under both daytime and nighttime traffic scenes. However, this technique suffers highly computational cost. Huang *et al.* [10] proposed a method based on block-based contrast analysis and inter-frame change information. This contrast-based method can effectively detect outdoor objects in a given surveillance area using a stationary camera. However, the contrast and inter-frame change information are sensitive to lighting effect caused by headlight beams of moving vehicles, and result in many erroneous results on vehicle detection.

In this study, we propose an effective nighttime vehicle detection and tracking approach for identifying and classifying moving vehicles by locating and analyzing spatial and temporal features of their vehicle lights for traffic surveillance. This proposed system comprises the following processing stages. First, a fast bright-object segmentation process based on automatic multilevel histogram thresholding is performed to extract pixels of lighting objects from the captured image sequences of nighttime traffic scenes. These lighting objects are then grouped by a spatial clustering process to obtain groups of vehicle lights of potential moving cars and motorbikes. Next, a feature-based vehicle tracking and identification process is applied to analyze the spatial and temporal information from these potential vehicle light groups from consecutive frames, and to refine the detection results and correct for grouping errors and occlusions. Actual vehicles and their types can thus be efficiently detected and verified from these tracked potential vehicles to obtain the traffic flow information in the road scenes. Experimental results demonstrate that the proposed approach is feasible and effective for vehicle detection and identification in various nighttime environments for traffic surveillance.

II. LIGHTING OBJECT EXTRACTION

The input image sequences grabbed from the vision system, which is mounted on the elevated platform of the highway and urban roads. The image sequences are grabbed with the 640x480 resolution with 24-bit true colors. Fig. 1 shows one sample nighttime traffic scene taken from the vision system. In this sample scene, there are moving cars and motorbikes on the road, and the salient features are their vehicle lights. In addition to the vehicle lights of the vehicles, some lamps, traffic lights and signs are also the visible illuminant appeared in the image sequences of the nighttime traffic scenes.

Fig. 1. An example of nighttime traffic scene

Hence, the first task is to extract these bright objects from the road scene image to facilitate further rule-based analysis. To save the computation cost on extracting bright objects, we firstly extracted the grayscale image, i.e. the Y-channel, of the grabbed image by performing a RGB to Y transformation. For extracting these bright objects from a given transformed gray-intensity image, pixels of bright objects must be separated from other object pixels of different illuminations. Thus, an effective multilevel thresholding technique is needed for automatically determining the appropriate number of thresholds for segmenting bright object regions from the traffic-scene image. For this purpose, we have already proposed an automatic multilevel thresholding technique for image segmentation [12]. This technique can automatically decompose a grabbed road-scene image into a set of homogeneous thresholded images. As shown in Fig. 2, bright objects of interest can be suitably extracted from the brightest one of the resultant thresholded images.

Fig. 2. Bright object plane extracted from Fig. 1 after performing the bright object segmentation process

III. SPATIAL CLUSTERING PROCESS

To obtain potential vehicle light components from the detection zone in the bright object plane, a connected-component extraction process [13] is performed to label and locate the connected-components of the bright objects. By extracting the connected- components, meaningful features of the location, dimension, and pixel distribution associated with each connected-component are obtained. The location and dimension of a connected-component can be represented by the bounding box which encloses it. Since there are various non-vehicle illuminant light components coexisted with actual vehicle lights, such as traffic lamps, road signs, road reflector plates, reflected beams, and some other illuminant objects, a spatial classification process is applied on these bright components to preliminarily detect potential vehicle lights and filter out non-vehicle components. Then these detected potential vehicle lights are processed by the following vehicle light tracking and identification process to obtain the actual moving vehicles.

First, the definitions used in the projection-based spatial classification process are described as follows,

1). C_i denotes one certain bright connected-component to be processed.

2). CG_k denotes a group of bright components, $CG_k = \{C_i, i=0,1,2,\ldots,p\}$, the total number of connected-components contained in CG_k is denoted as $N_{cc}(CG_k)$.

3). The location of the bounding boxes of a certain component C_i employed in the spatial clustering process are their top, bottom, left and right coordinates, and they are denoted as $t(C_i)$, $b(C_i)$, $l(C_i)$, and $r(C_i)$, respectively.

4). The width and height of a bright component C_i are denoted as $W(C_i)$ and $H(C_i)$, respectively.

5). The horizontal distance D_h and vertical distance D_v between two bright components are defined as,

$$D_h(C_i,C_j) = \max\left[l(C_i),l(C_j)\right] - \min\left[r(C_i),r(C_j)\right] \quad (1)$$

$$D_v(C_i,C_j) = \max\left[t(C_i),t(C_j)\right] - \min\left[b(C_i),b(C_j)\right] \quad (2)$$

If the two bright components are overlapping in the horizontal or vertical direction, then the value of the $D_h(C_i,C_j)$ or $D_v(C_i,C_j)$ will be a negative value.

6). Hence the measure of overlapping between horizontal and vertical projections of the two bright components can be respectively computed as,

$$P_v(C_i, C_j) = -D_v(C_i, C_j) / \min\left[H(C_i), H(C_j)\right] \quad (3)$$

$$P_v(C_i, C_j) = -D_v(C_i, C_j) / \min\left[H(C_i), H(C_j)\right] \quad (4)$$

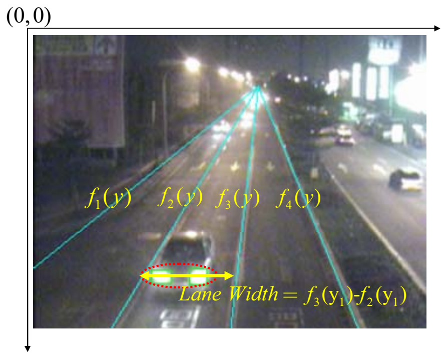

Fig. 3. The illustration of approximated lane widths on the image coordinate

An illustrative image coordinate system for vehicle detection is shown in Fig. 3. In the image coordinate system, the vehicles which are located at a distant place on the road will appear in the higher place and become progressively smaller, and will converge into a vanishing point. Therefore, the driving lanes stretched from the vanishing point can be modelized by a set of line equations by,

$$f_k(y) = \frac{y - c_k}{m_k} \quad ; \quad k = 1, 2, \ldots K \ , \quad (5)$$

where m_k and c_k are the slope and intercept of the *k*-th driving lane, respectively. Then the approximate lane width associated with a bright component C_i at a distance on the image coordinate, denoted by $LW(C_i)$, can be obtained by,

$$LW(C_i) = \left| f_{k+1}(C_Y(C_i)) - f_k(C_Y(C_i)) \right| \quad , \quad (6)$$

where $C_Y(C_i)$ represents the vertical position of the component C_i on the image coordinate, and is defined by $C_Y(C_i) = (t(C_i) + b(C_i))/2$. Based on the above-mentioned definitions of bright components, a classification procedure can be applied on these obtained bright components to preliminarily determine potential vehicle light components and filter out most non-vehicle illuminant light components.

For this purpose, a component C_i is determined as a potential vehicle light component if it satisfy the following conditions,

1). Sine a vehicle light mostly appear in nearly circular shape, the enclosing bounding box of a potential vehicle light component should form a square shape, i.e. the size-ratio feature of C_i must satisfy the following condition,

$$TH_{RL} \ \leq \ W(C_i)/H(C_i) \ \leq \ TH_{RH} \ , \quad (7)$$

where the thresholds TH_{RL} and TH_{RH} for the size-ratio condition are selected as 0.8 and 1.2, respectively, to suitably determine the circular-shaped appearance of a potential vehicle light.

2). Besides, a vehicle light object should have a reasonable area as compared to the one of the lane, thus the area feature of C_i must satisfy the following condition,

$$TH_{AL} < \ A(C_i) \ < TH_{AH} \quad (8)$$

where the thresholds TH_{AL} and TH_{AH} for the area condition are determined as $TH_{AL} = (LW(C_i)/8)^2$, and $TH_{AH} = (LW(C_i)/4)^2$, respectively, to adaptively reflect the reasonable area characteristics of a potential vehicle light.

After the preliminary classification procedure of bright components being performed, although most non-vehicle illuminant light components are filtered, there are still some vehicle-light-like components, such as reflected beams on the road ground. Therefore, a merging process is then applied to associate nearby potential vehicle light components which are belonged to the compound vehicle lights of the same car into light groups, and further filter out non-vehicle-light components with alike features.

Accordingly, if two neighboring bright components C_i and C_j satisfy the following conditions, they are determined as a homogeneous potential vehicle light set and are merged and clustered as the potential vehicle light set:

1). They are horizontally close to each other, i.e.:

$$D_h(C_i, C_j) \ < \ \min\left[W(C_i), W(C_j)\right], \quad (9)$$

2). They are also vertically close to each other, i.e.:

$$D_v(C_i, C_j) \ < \ 2.0 \times (\min\left[H(C_i), H(C_j)\right]), \quad (10)$$

3). The two vertically overlapping bright objects having high horizontal projection profiles should be grouped the same group *CG*:

$$P_h(C_i, C_j) \ > \ T_{hp} \quad (11)$$

where the threshold T_{hp} is chosen as 0.6, to respect the vertical alignment characteristics of compound vehicle lights.

These represent potential components of vehicles, as nominated as ***P*** in the following process. Notably, in the current stage, the vehicle light sets on the two sides of the vehicle body are not merged into pairing groups. This is because that there are vehicles with pairing light sets and motorbikes with single light sets existing in most nighttime road scenes, and it is not easy to determine whether nearing vehicle lights sets are pairing light sets belonged to the same vehicle without motion information in the subsequent frames. Thus, a vehicle light tracking and identification process described in the following section will be applied on these potential vehicle light sets to identify the actual moving vehicles and motorbikes.

IV. Tracking and Identification of Potential Vehicles and Motorbikes

Our proposed vehicle tracking and identification process includes three phases. First, the phase of potential vehicle component tracking process associates the motion relation of vehicle components in succeeding frames by analyzing their spatial and temporal features. Then the phase of motion-based grouping process is applied on the tracked vehicle components to form whole moving vehicles. Accordingly, these determined moving vehicles are then tracked by the vehicle tracking phase. Next, the phase of vehicle recognition process identifies and classifies the types of the tracked vehicles.

A. Tracking Process of Potential Vehicle Components

When a potential vehicle component is firstly detected in the field of view in front of the host car, a tracker will be created to associate this potential vehicle component with those in subsequent frames by applying their spatial-temporal features. The features used in the tracking process are described and defined as follows:

1). P_i^t denotes the i^{th} potential vehicle component appearing in the detection zone in frame *t*; and the location of P_i^t employed in the tracking process is represented by its central position, and can be expressed by,

$$P_i^t = \left(\frac{l\left(P_i^t\right)+r\left(P_i^t\right)}{2}, \frac{t\left(P_i^t\right)+b\left(P_i^t\right)}{2} \right) \tag{12}$$

2). The tracker TP_i^t is used to represent the trajectory of P_i which has been tracked in sequential frames 1 to *t*, and is defined as:

$$TP_i^t = \left\langle P_i^1, P_i^2, ..., P_i^t \right\rangle \tag{13}$$

3). The overlapping score of the two potential vehicle components P_i^t and $P_j^{t'}$, detected at two different time *t* and *t'*, can be computed by using their area of intersection,

$$S_o(P_i^t, P_j^{t'}) = \frac{A\left(P_i^t \cap P_j^{t'}\right)}{Max\left(A(P_i^t), A(P_j^{t'})\right)} \tag{14}$$

In each recursion of the tracking process for a newly incoming frame *t*, the potential vehicle components appearing in the incoming frame, denoted by $\mathbf{P}^t = \left\{ P_i^t \middle| \ i = 1,...,k' \right\}$, will be analyzed and associated with the set of potential vehicle components which have already been tracked in the previous frame *t*-1, denoted by $\mathbf{TP}^{t-1} = \left\{ TP_j^{t-1} \middle| \ j = 1,...,k \right\}$, and then the set of the tracked potential vehicles $\mathbf{TP}^t$ will be accordingly updated according to the following process.

During the tracking process, a potential vehicle component might be in one of three possible tracking states. The component tracking process applies different relevant operations according to the given states of each tracked potential vehicle component in each frame. The tracking states and associated operations for the tracked potential vehicle components are as follows:

1). **Update**: When a potential vehicle component $P_i^t \in \mathbf{P}^t$ in the current frame can match a tracked potential vehicle component $TP_j^{t-1} \in \mathbf{TP}^{t-1}$ if the following *matching condition* is satisfied, then the tracker updates the set of the tracked potential vehicle components $\mathbf{TP}^t$ by associating P_i^t with the tracker TP_j^t. The matching condition is determined as,

$$S_o\left(P_i^t, TP_j^{t-1}\right) \ > \ 0.25, \tag{15}$$

2). **Appear**: If a newly coming potential vehicle component $P_i^t \in \mathbf{P}^t$ cannot match any $TP_j^{t-1} \in \mathbf{TP}^{t-1}$ at the previous time, then a new tracker for this potential vehicle component is created and appended to the updated set $\mathbf{TP}^t$.

3). **Disappear**: A existing tracker of potential vehicle component $TP_j^{t-1} \in \mathbf{TP}^{t-1}$ cannot be matched by any newly coming potential vehicle components $P_i^t \in \mathbf{P}^t$. A tracked potential vehicle component may sometimes be temporarily sheltered or occluded in some frames, and will soon re-appear in subsequent frames, thus, to prevent such a potential vehicle component from being regarded as a newly appeared potential vehicle, its tracker will be retained in the subsequent five frames. If a tracker of potential vehicle component TP_j^{t-1} cannot be matched by any potential vehicles $P_i^t \in \mathbf{P}^t$ for more than five succeeding frames, then this potential vehicle component will be judged to have disappeared and its tracker will be removed from the tracker set $\mathbf{TP}^t$ in the following frames.

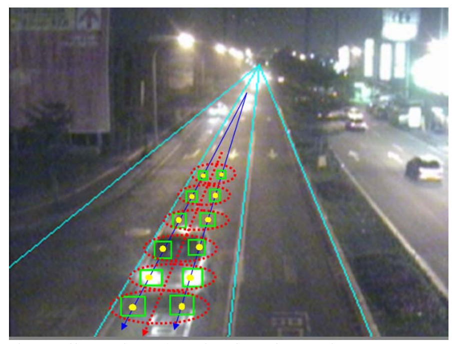

Fig. 4. Illustration of motion-based grouping process

B. Motion-Based Grouping of Vehicle Components

Having the tracks of potential vehicle components, the motion-based grouping process will then be applied to group together potential vehicle components which are belonged to the same vehicles. For this purpose, potential vehicle components which have rigidly similar motion in the successive frames will be grouped into a single vehicle. This concept is illustrated in Fig. 4.

Accordingly, pairing tracks of nearby potential vehicle components TP_i^t and TP_j^t are determined that they are belonged to the same vehicle if they have kept coherently moving and reveal homogeneous features for a period of time, and can be determined by the following coherent motion conditions,

1). They are consistently moving close with each other, i.e.:

$$\begin{aligned} &D_h(TP_i^{t-\tau}, TP_j^{t-\tau}) \ < \ \frac{LW(TP_i^{t-\tau}) + LW(TP_j^{t-\tau})}{2}, \\ &and \ \ D_v(TP_i^{t-\tau}, TP_j^{t-\tau}) \ < \ \frac{(\min\left[H(TP_i^{t-\tau}), H(TP_j^{t-\tau})\right])}{2} \end{aligned} \tag{16}$$

where $\tau = 0, \cdots, n-1$, n is determined to be 10, to appropriately reflect the sufficient sustained time of their common motion information.

2). They have similar heights for a span of time, i.e.:

$$H(TP_S^{t-\tau}) / H(TP_L^{t-\tau}) \; > \; T_h \qquad (17)$$

where $TP_S^{t-\tau}$ is the one with the smaller height among the two potential vehicle components $TP_i^{t-\tau}$ and $TP_j^{t-\tau}$ at the time $t-\tau$, while $TP_L^{t-\tau}$ is the larger one; and T_h is chosen to be 0.6 to reasonably reveal the alignment feature of pairing vehicle lights.

If the tracks TP_i^t and TP_j^t meet the above-mentioned coherent motion conditions, then they are merged into the same "component group track" of a potential vehicle, denoted by TG_k^t. After performing the motion-based grouping process on the vehicle component tracks, then a set of component group tracks, denoted by $\mathbf{TG}^t = \{TV_k^t \mid k = 1,...,k\}$, which consist of two or more vehicle components, can be obtained for the subsequent tracking process.

C. Tracking Process of Vehicle Component Groups

During a potential vehicles represented by a component group being tracked across the image, some possible occlusion problems caused by the segmentation process and the motion-based grouping process may occur. Therefore, having the potential vehicle tracks of component groups $TG_k^t \in \mathbf{TG}^t$ obtained by the motion-based grouping process, the component group tracking process will then be applied to accordingly update the position, motion and dimensions of each potential vehicle, and progressively refine the detection results of the potential vehicles by using spatial-temporal information in sequential frames. In this subsection, we present a tracking process for component groups of potential vehicles to handle the above-mentioned occlusion problems.

First, for each tracked component group of a potential vehicle till the previous frame $TG_k^{t-1} \in \mathbf{TG}^{t-1}$, its possible location in the current frame *t* will be preliminarily estimated by an adaptive search window based on its past motion information. To rapidly determine the search window of a tracked vehicle component group, its motion vector is firstly computed as,

$$\begin{aligned} \triangle x_k^{t-1} &= C_X(TG_k^{t-1}) - C_X(TG_k^{t-2}) \\ \triangle y_k^{t-1} &= C_Y(TG_k^{t-1}) - C_Y(TG_k^{t-2}) \end{aligned} \qquad (18)$$

, where $C_X(TG_k^t)$ and $C_Y(TG_k^t)$ respectively represents the horizontal and vertical position of tracked component group TG_k^t on the image coordinate, and are defined by $C_X(TG_k^t) = \left(l(TG_k^t) + r(TG_k^t)\right)/2$, and $C_Y(TG_k^t) = \left(t(TG_k^t) + b(TG_k^t)\right)/2$, respectively. Then a displacement factor (w_1, w_2) which reflects the potential position of the potential vehicle in the current frame can be computed as,

$$\begin{aligned} w_1 &= 1 + \frac{\triangle x_k^{t-1}}{\left\| \triangle x_k^{t-1}, \triangle y_k^{t-1} \right\|} \\ w_2 &= 1 + \frac{\triangle y_k^{t-1}}{\left\| \triangle x_k^{t-1}, \triangle y_k^{t-1} \right\|} \end{aligned} \qquad (19)$$

The center of the search window of a tracked potential vehicle in the current frame can be determined as $\left(w_1 \times C_X(TG_k^{t-1}), w_2 \times C_Y(TG_k^{t-1})\right)$, and its width and height can be defined as $1.5 \times W(TG_k^{t-1})$, and $3 \times H(TG_k^{t-1})$, respectively.

The possible positions of tracked potential components TP_i^t which are matched with a tracked potential component group TG_k^t in the current frame can be more rapidly and correctly obtained in the search window. For a tracked component group TG_k^t found in the search window, it may be in four possible states associated with its own component tracks TP_i^t, …, TP_{i+n}^t. This potential vehicle tracking process will conduct different operations according to the current state of TG_k^t as:

1). **Update**: all of the grouped component tracks TP_i^{t-1},…, TP_{i+n}^{t-1} owned by a tracked component group TG_k^{t-1} in the previous frame can still exactly and respectively match a set of the vehicle component tracks $TP_{i'}^t$, …, $TP_{i'+n}^t$ in the current frame within the search window, i.e. they all satisfy the following *matching condition*,

$$S_o\left(TP_{i'}^t, TG_k^{t-1}\right) \; > \; 0.5 \qquad (20)$$

Then the vehicle tracker just updates that the component group TG_k^t of a potential vehicle to be comprised of the renewed group of $TP_{i'}^t$, …, $TP_{i'+n}^t$.

2). **Shelter/Absorb**: the grouped component tracks TP_i^{t-1},…, TP_{i+n}^{t-1} owned by TG_k^{t-1} in the previous frame now have lesser number of component tracks $TP_{i'}^t$, …, $TP_{i'+m}^t$ (where $m < n$) to be found in the current frame within the search window. The matching condition (Eq. (20)) of the component group TG_k^{t-1} with $TP_{i'}^t$, …, $TP_{i'+m}^t$ will be respectively checked, then the ones which are satisfied with the matching condition will still be associated with the renewed TG_k^t. For the tracks of unexpectedly disappeared or absorbed components which are missed from TG_k^t, they will be still retained in the TG_k^t until they are regarded as disappeared components and removed by the potential vehicle component tracking process.

3). **Extend/Split**: the grouped component tracks TP_i^{t-1},…, TP_{i+n}^{t-1} owned by TG_k^{t-1} in the previous frame now are extend or split into more number of component tracks $TP_{i'}^t$, …, $TP_{i'+m}^t$ (where $m > n$) in the current frame within the search window. The matching condition (Eq. (20)) of TG_k^{t-1} with $TP_{i'}^t$, …, $TP_{i'+m}^t$ will be respectively checked, then the ones which are coincide with TG_k^{t-1} will still be associated with the renewed TG_k^t. For the tracks of newly appeared or split components which are not matched with TG_k^{t-1}, the motion-based grouping process (Eq. (16), and (17)) will be applied on these non-matched component tracks to check if they have coherent motion property with TG_k^{t-1}, and then the ones having coherent motion will be assigned into the renewed TG_k^t, and the others will be detached to be orphan component tracks.

4). **Exit**: a tracked potential component group TG_k^t has moved across the boundary of the detection region, and now all its own component tracks are determined to be disappeared by the potential vehicle component tracking process.

D. Vehicle Identification and Classification from Tracking Process

During the tracking process of the potential vehicles, a rule-based vehicle verification and classification process is then applied on each of the potential components and component groups of potential vehicles having been tracked for a time of more than 10 frames, to determine and classify whether it comprises a car, a motorcycle, or other on-road illuminated objects.

● **Car identification**

First, to identify moving cars in frame sequence, we can reasonably assume that a group of lighting components may have a higher possibility to be a car. Therefore, for a tracked component groups TG_k^t which has been consistently tracked by the component group tracking process for a span of more than 10 frames after being created by the motion-based grouping process, then TG_k^t can be nominated as a candidate of a moving car. Accordingly, if TG_k^t contains a set of actual vehicle lights that reveal an actual car, then TG_k^t mostly satisfy the following discriminating rules of statistical features:

1). Since a moving car can be approximately modeled as a rectangular patch, the enclosing bounding box of a potential car should form a horizontal rectangular shape, i.e. the size-ratio feature of the enclosing bounding box of TG_k^t must satisfy the following condition,

$$\tau_{r1} \leq W\left(TG_k^t\right)\big/H\left(TG_k^t\right) \leq \tau_{r2} \tag{21}$$

where the threshold τ_{r1} and τ_{r2} on the size-ratio condition are selected as 2.0 and 10.0 to suitably identify the rectangular-shaped appearance of paired vehicle lights.

2). Moreover, the number of the lighting components of TG_k^t should also be symmetrical and well-aligned, and thus the number of these components should be in reasonable proportion to the size of the size-ratio feature of its enclosing bounding box, thus, the following alignment condition should be satisfied,

$$\tau_{a1}\left(\frac{W(TP_j)}{H(TP_j)}\right) \leq N_{cc}\left(TP_j\right) \leq \tau_{a2}\left(\frac{W(TP_j)}{H(TP_j)}\right) \tag{22}$$

where the thresholds τ_{a1} and τ_{a2} are determined as 0.4 and 2.0, respectively, according to our analysis of typical visual characteristics of most moving cars appeared in the traffic scenes during nighttime driving.

● **Motorbike identification**

For the purpose of identifying the motorbikes, we can adopt the fact that a motorbike in the nighttime traffic scenes is mostly appeared as a single, and nearly square-shaped or vertical rectangular-shaped lighting component. Thus, for a single tracked component TP_i^t which has not been associated to any component groups and been consistently and alone tracked by the vehicle component tracking process for a span of more than 25 frames, then TP_i^t can be determined as a candidate of a moving motorbike. Therefore, if a single tracked component TP_i^t is actually a motorbike, then the size-ratio feature of its enclosing bounding box should reflect a square or vertical rectangular shape, and should satisfy the following discriminating rule:

$$\tau_{m1} \leq W\left(TP_i^t\right)\big/H\left(TP_i^t\right) \leq \tau_{m2} \tag{23}$$

where the threshold τ_{m1} and τ_{m2} on the size-ratio condition are selected as 0.6 and 1.2 to suitably identify the shape appearance characteristic of the motorbikes. Accordingly, a tracked component group or single potential component of a potential vehicle will be identified and classified as an actual car or a motorbike according to the above-mentioned vehicle classification rules.

V. Experimental Results

The proposed system is implemented on a Pentium-4 2.4 GHz platform which is set up on our elevated platform of the highway and urban roads. The frame rate of the vision system is 30 frames per second and the size of each frame of grabbed image sequences is 640 pixels by 480 pixels per frame.

TABLE 1. Experimental data of our proposed approach on an urban road

Lane	Detected Vehicles	Actual Vehicles
Lane 1	921	969
Lane 2	292	300
Lane 3	228	233
Total No. Cars	887	909
Total No. Motorbikes	584	593
Detection Rate of Cars	97.58%	
Detection Rate of Motorbikes	98.48%	
Time span of the video	50 minutes	

The proposed system has been tested on several videos of real nighttime traffic scenes in various conditions. Figures 5 – 6 exhibit the most representative ones of the experimental samples on performance evaluation. As shown in Fig. 5, the moving cars and motorbikes in an urban road are correctly detected and tracked by locating its vehicle lights, although some other non-vehicle illuminated objects also coexist with the vehicle in this scene. TABLE 1 depicts the quantitative results of the proposed approach on vehicle detection and tracking in urban road.

Fig. 5. Results of vehicle detection and tracking on the nighttime urban traffic scene

TABLE 2. Experimental data of our proposed approach on the highway scene

Lane	Detected Vehicles	Actual Vehicles
Lane 1	1392	1428
Lane 2	1527	1535
Lane 3	1495	1536
Total No. Cars	4397	4499
Detection Rate of Cars	97.73%	
Time span of the video	50 minutes	

Fig. 6. Results of vehicle detection and tracking on the nighttime highway traffic scene

As shown in Fig. 6, another sample of traffic scene of nighttime highway is illustrated. The vehicle lights of multiple moving vehicles are close to each other in this traffic scene, and the proposed method still successfully detect and track almost all vehicles by locating their vehicle light pairs.

TABLE 2 shows the quantitative results of the proposed approach on vehicle detection on the nighttime highway.

For an input video sequence with 640 x 480 pixels per frame, the proposed system takes an average of 16 milliseconds processing time per frame. This frugal computation cost ensures that the proposed system can effectively satisfy the demand of real-time processing at more than 30 frames per second. As can be seen from the experimental results, the proposed system demonstrates that it can provide fast, real-time, and effective nighttime vehicle detection and identification performance for traffic surveillance.

ACKNOWLEDGEMENTS

This work was supported by the National Science Council of R.O.C. under Contract No.: NSC- 97-2218-E-468-007.

REFERENCES

[1] S. Kamijo, Y. Matsushita, K. Ikeuchi, M. Sakauchi, "Traffic monitoring and accident detection at intersections", *IEEE Trans. Intell. Transport. Syst.*, vol. 1, no. 2, pp.108-118, 2000.

[2] A. H. S. Lai and N. H. C. Yung, "Vehicle-type identification through automated virtual loop assignment and block-based direction-biased motion estimation", *IEEE Trans. Intell. Transport. Syst.*, vol. 1, no. 2, pp.86-97, 2000.

[3] R. Cucchiara, M. Piccardi, and P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system," *IEEE Trans. Intell. Transport. Syst.*, vol. 1, no. 2, pp.119-130, Jun. 2000.

[4] M.-C. Huang and S.-H. Yen, "A real-time and color-based computer vision for traffic monitoring system", *IEEE Int'l Conf. Multimedia Expo*, pp.2119-2122, May. 2004.

[5] J. Kong, Y. Zheng, Y. Lu, B. Zhang, "A novel background extraction and updating algorithm for vehicle detection and tracking", *IEEE Int'l Conf. Fuzzy Syst. knowled. Discov.*, 2007.

[6] S. Gupte, O. Masoud, R. F. K. Martin, N. P. Papanikolopoulos, *IEEE Trans. Intell. Transport. Syst.*, vol. 3, pp. 37-47, 2002.

[7] B.-F. Wu, S.-P. Lin, Y.-H. Chen, "A Real-time multiple-vehicle detection and tracking system with prior occlusion detection and resolution", *IEEE Int'l Symp. Sig. Process. Info. Tech.*, pp.311-316, Dec. 2005.

[8] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters", *IEEE Conf. Comput. Vis. Patt. Recog.*, pp.495-510, Jun.1997.

[9] J. Zhou, D. Gao, and D. Zhang, "Moving vehicle detection for automatic traffic monitoring", *IEEE Trans. Veh. Tech.*, vol. 56, no. 1, pp.51-59, Jan. 2007.

[10] K. Huang, L. Wang, T. Tan, and S. Maybank, "A real-time object detecting and tracking system for outdoor night surveillance", *Pattern Recognit.*, vol. 41, pp.432-444, 2008.

[11] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Trans. Sys., Man, Cybern., vol. SMC-9, pp. 62-66, 1979.

[12] B.-F. Wu, Y.-L. Chen, and C.-C. Chiu, "A discriminant analysis based recursive automatic thresholding approach for image segmentation," IEICE Trans. Info. Systems, vol. E88-D, no.7, pp.1716-1723, 2005.

[13] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations", Computer Vision & Image Understand., vol. 89, pp. 1-23, 2003.